

A process for sequence saturation mutagenesis (SeSaM)

Introduction:

5 Inspired by Darwinian evolution in nature, random mutagenesis methods have been developed for tailoring proteins to our needs and elucidating structure-function relationships (1). Creating diversity on the genetic level is the complementary tool to gene shuffling since it creates the opportunity to introduce novel mutations for adapting proteins to non-natural environments in biotechnological processes. The sequence space 10 of a truly randomized library is, by its nature, not limited by a pre-selection for function under physiological conditions. Parental genes used in gene shuffling experiments are optimized by natural evolution for function under physiological conditions.

15 Among random mutagenesis methods, error-prone PCR methods based on inaccurate amplification of genes are most commonly used due to their simplicity and versatility. Error-prone PCR methods can be divided into three categories: A) Methods that reduce the fidelity of the polymerase by unbalancing nucleotides concentration and/or adding of manganese chloride (2-4), B) Methods that employ nucleotide analogs (5,6), and C) Combined methods (A and B; (7)).

20

Specification

The invention relates to a process for the mutagenesis of a double-stranded polynucleotide sequence (master sequence) of n base-bairs having a (+)-strand and a complementary (-)-strand comprising the steps

- (i) creation of a collection of single-stranded fragments of the (+)-strand of the master sequence wherein all members of the collection have the same 5'-terminus and have a deletion in the 3'-terminus such that the collection represents (+)-strands with a length of $n-1, n-2, n-3, \dots$ nucleotides;
- 30 (ii) introduction of at least one universal nucleotide or degenerate nucleotide at the 3'-terminus of the (+) strand produced in step (i);
- (iii) elongation of the (+)-strand produced in step (ii) to the full length of the master sequence using the (-)-strand or fragments thereof as a template strand for the 35 elongation;
- (iv) synthesis of a (-)-strand by using the (+)-strand produced in step (iii) as a template strand thereby effecting mutations in the (-)-strand at the positions of the previous universal nucleotides or degenerate nucleotides compared to the master sequence.

40

A universal nucleotide is a nucleotide that can pair to all four standard nucleotides. For example deoxyinosinetriphosphate (dITP) when incorporated into a polynucleotide strand allows base pairing with Adenine, Guanine, Thymine and Cytosine.

5 Preferred universal nucleotides are deoxyinosine, 3-nitropyrrole and 5-nitroindole.

A degenerate nucleotide is a nucleotide that can pair to less than all four standard nucleotides. Preferred degenerate nucleotides are N⁶-methoxy-2,6-diaminopurine (K), N⁶-methoxy-aminopurine (Z), hydroxylaminopurine (HAP), 2'-deoxyribonucleoside triphosphate (dTP), 6H,8H-3,4-dihydropyrimidol [4,5-c][1,2] oxazin-7-one (P), N⁴-aminocytidine, N⁴-hydroxy-2'-deoxycytidine, N⁴-methoxy-2'-deoxycytidine and 8-oxodeoxyguanosine triphosphate (8-oxo-G).

10 A nucleotide analog with promiscuous base pairing property means that a nucleotide that is based on a purine or pyrimidine structure (e.g. A,G,C,T) which is modified in one or more of its functional groups, can have base pairs to more than one other nucleotides, e.g. to two, three or four nucleotides. Also the universal nucleotides and degenerate nucleotides are embraced by the term nucleotide analog with promiscuous base pairing property.

15 20 A preferred embodiment of the invention is a process as described above, wherein an oligonucleotide of the general formula

p(U)_a(N)_b*(S)_c[TERM]

25 with
p = 5'-phosphate or hydroxy-group or any chemical group capable of forming diester bonds
U = universal or degenerate bases
a = arbitrary integral number from 0 to 10000, preferred from 1-100
30 N = mixture of four bases (A/T/G/C (standard nucleotides))
b = arbitrary integral number from 0 to 100, preferred from 1-10,
* = cleavable group such as phosphothioate bonds in phosphothioate nucleotides
S = standard nucleotide or nucleotide analog
35 c = arbitrary integral number from 0 to 100, preferred from 1-10,
[TERM] = a dye terminator or any group preventing elongation of the oligonucleotide, with the proviso that a+b>0, preferred >1, more preferred >2, is used in step (ii) to introduce universal or degenerate bases to the collection of single-stranded fragments created in step (i).

40 [TERM] can be any group which prevents the elongation of the above-mentioned oligonucleotide, preferably an hydrogen or a dye terminator. Preferred dye terminators

are Coumarin, 6-FAM, Fluorescein, Fluorescein-dT, JOE, Oregon Green, ROX, TAMRA or Texas Red-X.

Another preferred embodiment of the invention is the creation of single-stranded fragments of the (+)-strand of the master sequence according to step (i) by incorporating alpha-phosphothioate nucleotides, preferably dATP α S, dGTP α S, dTTP α S, dCTP α S, into the PCR products and subsequent cleavage of the phosphothioate bond by iodine under alkaline conditions.

10 Another preferred embodiment of the process according to the invention is the following process for step (iii): Starting from a double-stranded plasmid which harbors the master sequence a (-) single stranded plasmid polynucleotide sequence is synthesized using a primer which anneals downstream of the (+)-strand of the master sequence. This (-) single stranded plasmid polynucleotide sequence is annealed with the 15 (+)-strands produced in step (ii). The (+)-strands are elongated to the full length of the master sequence using the (-)-strand as a template. This process is shown in Figure 3.

Further preferred embodiments are disclosed in the claims.

20 Step (i): Creating DNA fragment pool with length distribution

In the first step, PCR is performed using biotinylated forward primer and non-biotinylated reverse primer in the presence of both standard nucleotides and α -phosphothioate nucleotides. α -Phosphothioate nucleotides are similar to normal nucleotides, except that an oxygen atom of α -phosphate is replaced by a sulphur atom. 25 Phosphothioate bond is susceptible to iodine cleavage in alkaline condition. Due to random distribution of phosphothioates in the DNA, a library of fragments that stop at every single base is generated in a single PCR. Nicks may form if the binding between two complementary strands is strong. Biotinylated fragments are isolated using Streptavidin-coated biomagnetic beads. DNA melting solution (0.1 M NaOH) is then used to 30 remove non-biotinylated strands and undesired fragments. Biotinylated fragments can easily be released from biomagnetic beads by boiling in 0.1 % SDS solution. A scheme of the first step is shown in Figure 1a and the corresponding experimental data is shown in Figure 1b.

35

Step (ii): Enzymatic elongation

To elongate DNA fragments with universal bases or degenerate bases (Figure 2), two approaches can be used. In a first approach terminal deoxynucleotidyl transferase has 40 been used for incorporating ambiguous bases (universal base or degenerate base) at the 3'-termini. The second approach requires single-stranded DNA ligation between DNA fragments and a "special" oligo. This "special" oligo is 5'-phosphorylated to facili-

tate ligation. It is terminated with fluorescein to avoid intra- and intermolecular ligation and to quantify the incorporation. There are 3 distinct parts in this oligo: 1) "Mutational part" containing universal bases or degenerate bases, 2) "Adhesive part" consisting of three bases that encompasses all 64 possibilities by using equimolar of A/T/G/C in 5 oligo synthesis. This part is designed to assist annealing in the subsequent PCR used for the full length gene synthesis. The "Redundant part" is connected to "Adhesive part" via phosphothioate bond that allows its cleavage by the iodine method. ssDNA ligation can be accomplished using ThermoPhage RNA Ligase II (Prokaria). ThermoPhage RNA Ligase II catalyses the ATP-dependent intra- and intermolecular formation of 10 phosphodiester bonds between 5'-phosphate and 3'-hydroxyl termini of single-stranded DNA or RNA. This enzyme is derived from thermophilic phage TS2126 that infects the thermophilic eubacterium *Thermus scotoductus*. This thermostable enzyme is homologous to RNA ligase derived from bacteriophage T4. It shows superior efficiency in ssDNA ligation as compared to T4 RNA ligase. The ligation efficiency was 15 determined by ligating a fluorescein labelled oligo to single stranded DNA template. After the ligation, the "Redundant part" is removed by iodine cleavage in alkaline condition.

Step (iii): Full-length gene synthesis

20 The third step is extending the elongated fragments to full-length (Figure 3). Here, we use a single-stranded template to avoid the wild-type amplification. Single-stranded template is synthesized using reverse primer. Methylated and hemimethylated parental genes are removed by *Dpn* I digestion. This procedure is similar to QuikChange Site-Directed Mutagenesis (Stratagene) except that only one non-mutagenic primer is 25 used instead of a pair of mutagenic primers. Elongated fragments anneal to single-stranded template due to complementarities and extend the single strand to full-length. Reverse primers present in this PCR reaction can only anneal to the newly synthesized 30 full-length single strand and not to single-stranded template. After reverse primer binding the double stranded full-length genes will be synthesized. Double-stranded DNAs will contain nucleotide analogs in one strand and standard nucleotides in another strand.

35 In the event where a double-stranded template is used in this PCR instead of a single-stranded one we would observe that the reverse primer binds to its complementary template strand, amplifying double stranded DNA that do not contain nucleotide analogs.

Step(iv): Nucleotide replacement

40 In the last PCR the nucleotide analog-containing strands are used as templates to replace nucleotide analogs with standard nucleotides (Figure 4). After restrictive diges-

tion, mutated genes are cloned into suitable expression vector, transformed and expressed in *E. coli*. Randomly selected clones were picked and grown in small cultivation tubes (5 ml; LB_{amp}). Isolated plasmid DNA is subsequently sequenced. Preliminary sequencing results with 100 clones proved the proper replacement and showed a bias 5 as expected for inosine (Figure 5 and Figure 6).

Generation of mutant libraries using the SeSaM method can be completed within 1-2 days. The SeSaM method offers the following advantages:

- 10 1) Able to saturate every single position of a sequence with all 20 possible naturally occurring amino acids.
- 2) No bias in mutational spectra if truly universal bases are used.
- 3) Mutation spectra can be manipulated using transition-favoured or transversion-favoured degenerate bases.
- 15 4) Controllable length of mutation regions by designing proper special oligos.
- 5) Fragment size distribution can be controlled by using Sp-dATP α S/ Sp-dTTP α S/ Sp-dGTP α S/ Sp-dCTP α S or combination of them.

Materials and Methods

20 All chemicals used were of analytical-reagent grade or higher quality and were purchased from Sigma-Aldrich Chemie GmbH (Taufkirchen, Germany), Applichem GmbH (Darmstadt, Germany) or Carl Roth GmbH +Co (Karlsruhe, Germany). pEGFP plasmid was purchased from BD Biosciences (Heidelberg, Germany).

25 A thermocycler (Mastercycler gradient; Eppendorf, Hamburg, Germany) and thin-wall PCR tubes (Multi-Ultra tubes; 0.2 ml; Carl Roth GmbH +Co., Karlsruhe, Germany) were used in all PCRs. The reaction volume of all PCRs was always 50 μ l.

30 **(Preparatory step 1) Single-stranded pEGFP preparation:**
For each PCR, 5 U Pfu Turbo polymerase (Stratagene, Amsterdam, Netherlands), 0.2 mM dNTP mix (New England Biolab, Frankfurt, Germany), 12.6 pmol reverse primer (5'-GACCGGCGCTCAGTTGGAATTCTAG-3') and 48.6-54.3 ng plasmid pEGFP (Miniprep, Qiagen, Hilden, Germany) were used. After PCR (95°C for 30 sec 35 1 cycle, 95°C for 30 sec/55°C for 1 min/68°C for 4 min 40 cycles), 40 U of *Dpn* I was added followed by incubation at 37°C for 3 hours. Product recovery was done using a NucleoSpin Extract (Macherey-Nagel, Düren, Germany; elution volume of 35 μ l for 200 μ l PCR product).

40 **(Preparatory step 2) Cloning vector preparation:**
24.3-27.1 μ g plasmid pEGFP (Miniprep; QIAGEN) was first digested with 30 U *Eco*RI (New England Biolab) in a reaction mixture of 100 μ l. The reaction mixture was incu-

bated at 37°C for 3 hours. The linearized plasmid was purified with NucleoSpin Extract (Machery-Nagel; Elution volume of 50 µl for 100 µl reaction mixture). 4.9-5.6 µg of linearized plasmid pEGFP was subjected to second digestion with 20 U Age I (New England Biolab) in a reaction volume of 50 µl. After incubation at 37°C for 3 hours, the 5 double digested plasmid was purified with NucleoSpin Extract (Macherey-Nagel; elution volume of 35 µl for 50 µl reaction mixture). Double digested pEGFP was used for subsequent cloning.

(Step 1) PCR with dATP α S:

10 For each PCR (94°C for 3 min 1 cycle, 94°C for 1 min/59.5°C for 1 min/72°C for 75 sec 31 cycles, 72°C for 10 min 1 cycle), 2.5 U Taq DNA polymerase (Qiagen), 0.2 mM dNTP mix (New England Biolab), 0.2 mM Sp-dATP α S (Biolog Life Science Institute, Bremen, Germany), 12.6 pmol 5'-biotinylated forward primer (5'-GACCATGATTACGCCAAGCTTGC-3'), 12.6 pmol reverse primer (5'-GAC 15 CGCGCTCAGTTGGAATTCTAG-3') and 242.9-271.4 ng plasmid pEGFP (Miniprep; Qiagen) were used.

(Step 2) Iodine cleavage of thiophosphodiester backbone:

20 The phosphothiate bond was cleaved with iodine (dissolved in ethanol; final concentration in the PCR tube 2 µM). The mixture was incubated at room temperature for 1 hour.

(Step 3) Preparation of single stranded DNA fragments with different length:

Biotinylated DNA fragments from the forward primer were isolated using Dynabeads MyOne Streptavidin (DYNAL Biotech, Oslo, Norway) at room temperature. 50 µl of bi- 25 omagnetic beads (10 mg/ml) were washed twice using 100 µl 2x B&W buffer (10 mM Tris-HCl, pH 7.5; 1.0 mM EDTA; 2.0 M NaCl). Washed biomagnetic beads were resuspended in 100 µl 2x B&W buffer and 100 µl of cleaved PCR product was added. After incubating for 20 min, the biotinylated DNA fragments were immobilized on the bio- magnetic beads and the iodine was removed by washing with 100 µl 2x B&W buffer. 30 Non-biotinylated DNA fragments were released after incubating in 100 µl DNA melting solution (0.1 M NaOH) at 37°C for 10 min followed by bead washing steps with 100 µl of DNA melting solution and 100 µl 1x B&W buffer. The washed Dynabeads were boiled in 60 µl 0.1 % SDS for releasing the DNA fragments from the solid support. The supernatant containing the DNA fragments was transferred to another tube immedi- 35 ately.

(Step 4) SDS salt removal from eluted single stranded DNA fragments:

Desalting was done using NucleoTrap kit (Macherey-Nagel, Düren, Germany). 400 µl of buffer NT2 was added to 100 µl of eluted DNA followed by 15 µl of NucleoTrap sus- 40 pension. The mixture was incubated at room temperature for 10 min and gently shaked every 2-3 min. The sample was then centrifuged at 10000 g for 30 sec and after dis-

carding the supernatant, the beads were washed with 500 µl of buffer NT3. The latter step was repeated once. The pellet was air-dried for 15 min at 37°C to remove residual ethanol, resuspended in 55 µl Tris/HCl buffer (5 mM; pH 8.5) and incubated for DNA elution at 50°C for 5 min for DNA elution. The suspension was pipetted into NucleoSpin

5 Microfilter and centrifuged at 10000 g for 30 sec to separate beads from DNA containing solution.

(Step 5) Enzymatic elongation of DNA fragments with universal base:

Total reaction volume for each elongation reaction was 50 µl. In each reaction, 5 U

10 terminal transferase (New England Biolabs), 0.25 mM CoCl₂, 0.4 µM dITP (Amersham Biosciences Europe GmbH, Freiburg, Germany), and 18 µl desalted DNA from step 4 were used. After incorporation of universal bases in the elongation reaction (37°C for 30 min and heat deactivation of the transferase at 70 °C for 10 min), the product was purified following the QIAquick Nucleotide Removal Kit (QIAGEN; elution volume of 25

15 µl for 50 µl reaction mixture) protocol.

(Step 6) Full-length gene synthesis:

For each PCR (94°C for 3 min 1 cycle, 94°C for 1 min/59.5°C + 0.2°C (0.2°C increment for each cycle) for 1 min/72°C for 3 min 30 cycles, 72°C for 10 min 1 cycle), 2.5 U Taq

20 DNA polymerase (Qiagen), 0.2 mM dNTP mix (New England Biolab), 13.3 µl elongated DNA fragment, 20 pmol reverse primer (5'-GAC CGGCGCTCAGTTGGAATTCTAG-3') and 0.66-0.76 µg single-stranded reverse template (preparatory step 1) were used. After synthesizing the full-length gene a purification step was performed using the NucleoSpin Extract (Macherey-Nagel; elution volume of 35 µl for 150 µl PCR product).

25

(Step 7) Universal base replacement:

For each PCR (94°C for 3 min 1 cycle, 94°C for 1 min/52.7°C for 1 min/72°C for 75 sec 30 cycles, 72°C for 10 min 1 cycle), 2.5 U Taq DNA polymerase (Qiagen), 0.2 mM dNTP mix (New England Biolabs), 20 pmol forward primer (5'-

30 GACCATGATTACGCCAAGCTTGC-3'), 20 pmol reverse primer (5'-GAC CGGCGCTCAGTTGGAATTCTAG-3') and 2.5 µl full-length gene (step 6) were used. The PCR product was purified using NucleoSpin Extract (Macherey-Nagel; elution volume of 50 µl for 150 µl PCR product).

35

(Step 8) PCR product digestion:

After universal base replacement, 40 µl of purified PCR product (step 7) was digested with 30 U EcoRI (New England Biolab) for 3 hours at 37°C. 20 U of Agel (New England Biolab) was then added to the digest and the mixture was incubated for additional 3

40 hours at 37°C. The digested product was purified using NucleoSpin Extract (Macherey-Nagel; elution volume of 25 µl).

(Step 9) Ligation and Transformation:

The ligation of the digested PCR product (step 8) and pEGFP cloning vector (preparatory step 2) was performed at room temperature for 1 hour using T4 DNA ligase

5 (Roche, Mannheim, Germany) and transformed into *E. coli* XL2 Blue (Stratagene, Amsterdam, Netherlands) cells. Competent cells were prepared by resuspending the cell pellet of a 50 ml culture (OD_{578} 0.4-0.5) in 2 ml TSS buffer (10 g PEG 6000; 5 ml DMSO; 0.6 g $MgSO_4$; 100 ml LB). 5 μ l ligation mixture was added to 200 μ l cell aliquot followed by incubation in ice for 20 min, heat-shock s at 42 °C for 45 sec and additional chilling in ice for 2 min. After adding 0.8 ml of LB, the culture was shaken at 37°C and 170 rpm for 1 hour. Cells were harvested by centrifugation at 3000 g, room temperature for 2 min. 900 μ l of supernatant was discarded and cells were gently re-suspended in the remaining 100 μ l of supernatant. The cells were then plated on LB/Amp plate and incubated at 37°C for overnight.

15

Description of the Figures

Figure 1: a) Step (i): A random fragment size distribution is created b) Upper and middle gel picture: PCR product before (left lane) and after (right lane) iodine cleavage.

20 Lower gel picture: DNA fragment size distribution after DNA melting and purification for different concentration of Sp-dATP α S (stated at the lower part of the lane).

Figure 2: a) Step (ii): DNA fragments elongated with universal or degenerate bases b) Lower left gel picture: PCR with primers elongated with different concentration of deoxyinosine at 3'-termini using terminal deoxynucleotidyl transferase. Lower right gel picture: PCR with primers elongated with different concentration of 5-nitroindole at 3'-termini using terminal deoxynucleotidyl transferase.

25 Figure 3: Step (iii): Synthesis of full-length genes containing nucleotide analogs

30 Figure 4: Step (iv): Nucleotide analogs are replaced by standard nucleotides.

Figure 5: Sequencing result of 100 randomly picked clones.

35 Figure 6: Random distribution of mutations of 100 sequenced clones.

Literature

1. Arnold, F.H., Wintrode, P.L., Miyazaki, K. and Gershenson, A. (2001) *Trends Biochem. Sci.*, 26, 100-106.
2. Cadwell, R.C. and Joyce, G.F. (1994) *PCR Meth. App.*, 3, 136-140.

3. Lin-Goerke, J.L., Robbins, D.J. and Burczak, J.D. (1997) *Biotechniques*, 23, 409-412.
4. Cadwell, R.C. and Joyce, G.F. (1992) *PCR Meth. Appl.*, 2, 28-33.
5. Kuipers, O.P. (1996) *Meth. Mol. Biol.*, 57, 351-356.
6. Zaccolo, M., Williams, D.M., Brown, D.M. and Gherardi, E. (1996) *J. Mol. Biol.*, 255, 589-603.
7. Xu, H., Petersen, E.I., Petersen, S.B. and el-Gewely, M.R. (1999) *Biotechniques*, 27, 1102-1108.